3.3 THE GOLDBACH CONJECTURE

While there is an infinitude of primes, their distribution within the positive integers is most mystifying. Repeatedly in their distribution one finds hints or, as it were, shadows of a pattern; yet an actual pattern amenable to precise description remains unfound. The difference between consecutive primes can be small as with the pairs 11 and 13, 17 and 19, or for that matter 1,000,000,000,061 and 1,000,000,000,063. At the same time there exist arbitrarily long intervals in the sequence of integers which are totally devoid of any primes.

It is an unanswered question whether there are infinitely many pairs of *twin primes*; that is, pairs of successive odd integers p and $p+2$ which are both primes. Numerical evidence leads us to suspect an affirmative conclusion. Electronic computers have discovered 152,892 pairs of twin primes less than 30,000,000 and twenty pairs between 10^{12} and $10^{12}+10{,}000$, which hints at their growing scarcity as the positive integers increase in magnitude.

Consecutive primes can not only be close together, but also be far apart; that is, arbitrarily large gaps can occur between consecutive primes. Stated precisely: Given any positive integer n, there exist n consecutive integers, all of which are composite. To prove this, we need simply consider the integers

$$(n+1)!+2,\ (n+1)!+3,\ \ldots,\ (n+1)!+(n+1),$$

where $(n+1)!=(n+1)\cdot n\cdots 3\cdot 2\cdot 1$. Clearly there are n integers listed and they are consecutive. What is important is that each integer is composite; for, $(n+1)!+2$ is divisible by 2, $(n+1)!+3$ is divisible by 3, and so on.

For instance, if a sequence of four consecutive composite integers is desired, then the argument above produces 122, 123, 124, and 125:

$$\begin{aligned} 5!+2 &= 122 = 2\cdot 61,\\ 5!+3 &= 123 = 3\cdot 41,\\ 5!+4 &= 124 = 4\cdot 31,\\ 5!+5 &= 125 = 5\cdot 25. \end{aligned}$$

Of course, one can find other sets of four consecutive composites, such as 24, 25, 26, 27 or 32, 33, 34, 35.

This brings us to another unsolved problem concerning primes, the Goldbach Conjecture. In a letter to Euler (1742), Christian Goldbach hazarded the guess that every even integer is the sum of two numbers

that are either primes or 1. A somewhat more general formulation is that every even integer greater than 4 can be written as a sum of two odd prime numbers. This is easy to confirm for the first few even integers:

$$
\begin{aligned}
2 &= 1 + 1 \\
4 &= 2 + 2 = 1 + 3 \\
6 &= 3 + 3 = 1 + 5 \\
8 &= 3 + 5 = 1 + 7 \\
10 &= 3 + 7 = 5 + 5 \\
12 &= 5 + 7 = 1 + 11 \\
14 &= 3 + 11 = 7 + 7 = 1 + 13 \\
16 &= 3 + 13 = 5 + 11 \\
18 &= 5 + 13 = 7 + 11 = 1 + 17 \\
20 &= 3 + 17 = 7 + 13 = 1 + 19 \\
22 &= 3 + 19 = 5 + 17 = 11 + 11 \\
24 &= 5 + 19 = 7 + 17 = 11 + 13 = 1 + 23 \\
26 &= 3 + 23 = 7 + 19 = 13 + 13 \\
28 &= 5 + 23 = 11 + 17 \\
30 &= 7 + 23 = 11 + 19 = 13 + 17 = 1 + 29.
\end{aligned}
$$

It seems that Euler never tried to prove the result, but, writing to Goldbach at a later date he countered with a conjecture of his own: any even integer (≥ 6) of the form $4n + 2$ is a sum of two numbers each being either primes of the form $4n + 1$ or 1.

The numerical evidence for the truth of these conjectures is overwhelming (indeed Goldbach's Conjecture has been verified for all even integers up to 100,000), but a general proof or counterexample is still awaited. The nearest approach of modern number theorists to Goldbach's Conjecture is the result of the Russian mathematician Vinogradov, which states: Almost all even integers are the sum of two primes. The technical meaning of the term "almost all" is that if $A(n)$ denotes the number of even integers $m \leq n$ which are not representable as the sum of two primes, then

$$\lim_{n \to \infty} A(n)/n = 0.$$

As Landau so aptly put it, "The Goldbach conjecture is false for at most 0% of all even integers; this at most 0% does not exclude, of course, the possibility that there are infinitely many exceptions."

We remark that if the conjecture of Goldbach is true, then each odd number larger than 7 must be the sum of three odd primes. For, take n to be an odd integer greater than 7, so that $n - 3$ is even and greater

than 4; if $n-3$ could be expressed as the sum of two odd primes, then n would be the sum of three. In 1937, Vinogradov showed that this does indeed hold for every sufficiently large odd integer, say greater than N. Thus, it is enough to answer the question for every odd integer n in the range $9 \leq n \leq N$, which for a given integer becomes a matter of tedious computation (unfortunately, N is so large that this exceeds the capabilities of the most modern electronic computers).

Vinogradov's result implies that every sufficiently large even integer is the sum of not more than four odd primes. Thus, there is a number N such that every even integer beyond N is the sum of either two or four odd primes.

Having digressed somewhat, let us observe that according to the Division Algorithm, every positive integer can be written uniquely in one of the forms

$$4n,\ 4n+1,\ 4n+2,\ 4n+3$$

for some suitable $n \geq 0$. Clearly, the integers $4n$ and $4n+2=2(2n+1)$ are both even. Thus, all odd integers fall into two progressions: one containing integers of the form $4n+1$,

$$1, 5, 9, 13, 17, 21, \ldots$$

and the other containing integers of the form $4n+3$,

$$3, 7, 11, 15, 19, 23, \ldots .$$

While each of these progressions includes some obviously prime numbers, the question arises as to whether each of them contains infinitely many primes. This provides a pleasant opportunity for a repeat performance of Euclid's method for proving the existence of an infinitude of primes. A slight modification of his argument reveals that there are an infinite number of primes of the form $4n+3$. We approach the proof through a simple lemma.

LEMMA. *The product of two or more integers of the form* $4n+1$ *is of the same form.*

Proof: It is sufficient to consider the product of just two integers. Let $k=4n+1$ and $k'=4m+1$. Multiplying these together, we obtain

$$\begin{aligned} kk' &= (4n+1)(4m+1) \\ &= 16nm+4n+4m+1 = 4(4nm+n+m)+1, \end{aligned}$$

which is of the desired form.

This paves the way for:

THEOREM 3-6. *There is an infinite number of primes of the form* $4n + 3$.

Proof: In anticipation of a contradiction, let us assume that there exist only finitely many primes of the form $4n + 3$; call them $q_1, q_2, \ldots, q_s$. Consider the positive integer

$$N = 4q_1 q_2 \cdots q_s - 1 = 4(q_1 q_2 \cdots q_s - 1) + 3$$

and let $N = r_1 r_2 \cdots r_t$ be its prime factorization. Since N is an odd integer, we have $r_k \neq 2$ for all k, so that each r_k is either of the form $4n + 1$ or $4n + 3$. By the Lemma, the product of any number of primes of the form $4n + 1$ is again an integer of this type. For N to take the form $4n + 3$, as it clearly does, N must contain at least one prime factor r_i of the form $4n + 3$. But r_i cannot be found among the listing $q_1, q_2, \ldots, q_s$, for this would lead to the contradiction that $r_i \mid 1$. The only possible conclusion is that there are infinitely many primes of the form $4n + 3$.

Having just seen that there are infinitely many primes of the form $4n + 3$, one might reasonably ask: Is the number of primes of the form $4n + 1$ also infinite? This answer is likewise in the affirmative, but a demonstration must await the development of the necessary mathematical machinery. Both these results are special cases of a remarkable theorem by Dirichlet on primes in arithmetic progressions, established in 1837. The proof is much too difficult for inclusion here, so that we content ourselves with the mere statement.

THEOREM 3-7 (Dirichlet). *If a and b are relatively prime positive integers, then the arithmetic progression*

$$a, a + b, a + 2b, a + 3b, \ldots$$

contains infinitely many primes.

There is no arithmetic progression $a, a + b, a + 2b, \ldots$ that consists solely of prime numbers. To see this, suppose that $a + nb = p$, where p is a prime. If we put $n_k = n + kp$ for $k = 1, 2, 3, \ldots$, then the n_kth term in the progression is

$$a + n_k b = a + (n + kp)b = (a + nb) + kpb = p + kpb.$$

Since each term on the right-hand side is divisible by p, so is $a + n_k b$. In other words, the progression must contain infinitely many composite numbers.

It has been conjectured that there exist arithmetic progressions of finite (but otherwise arbitrary) length, composed of consecutive prime numbers. Examples of such progressions consisting of three and four